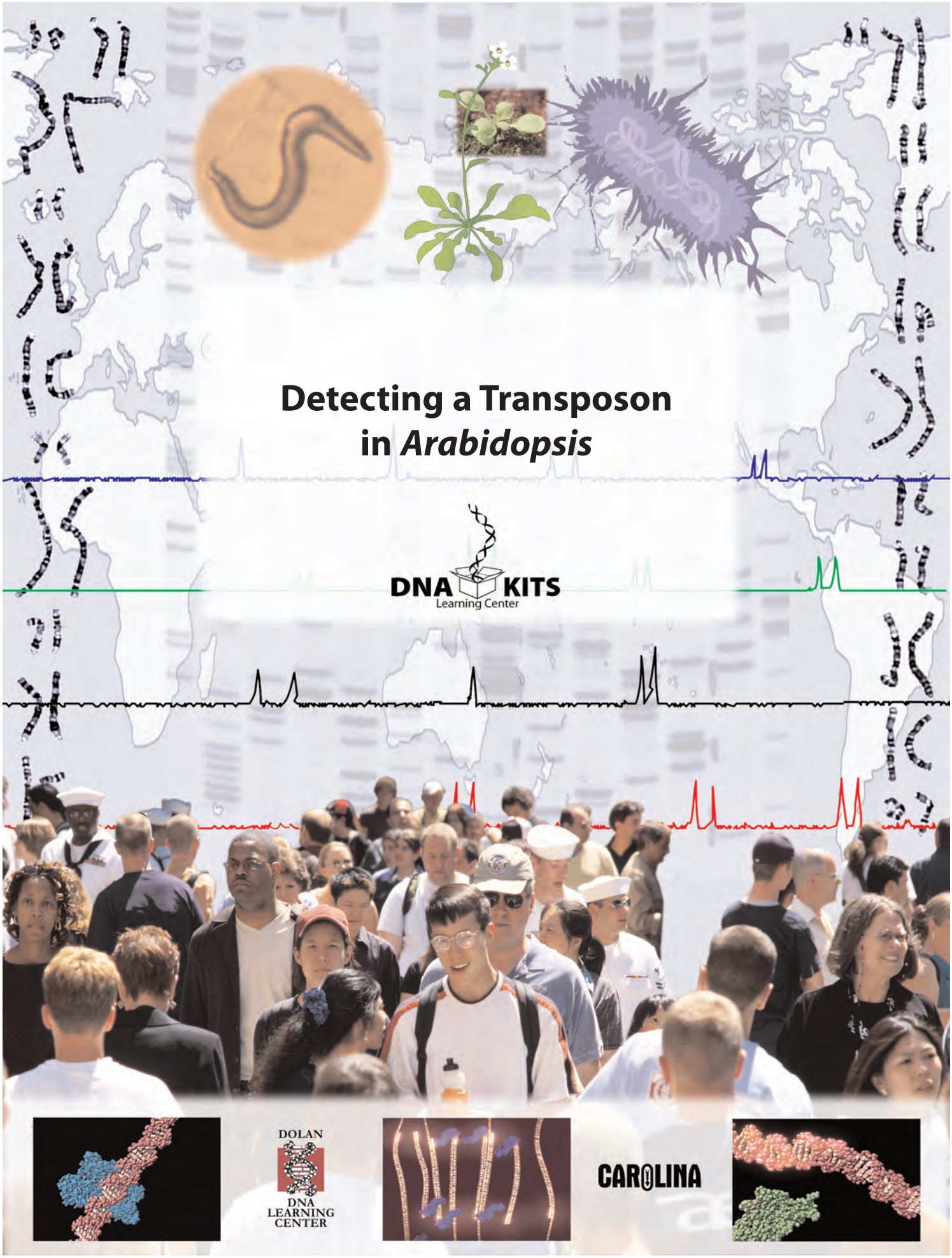


Detecting a Transposon in *Arabidopsis*

DNA KITS
Learning Center

DOLAN
DNA
LEARNING
CENTER

CAROLINA

Detecting a Transposon in *Arabidopsis*

IMPORTANT INFORMATION

Storage: Upon receipt of the kit, store primer/loading dye mixes, tris/EDTA (TE) buffer, and DNA marker pBR322/*Bst*NI in the freezer (approximately –20°C). All other materials may be stored at room temperature (approximately 25°C).

Use and Lab Safety: The materials supplied are for use with the method described in this kit only. Use of this kit presumes and requires prior knowledge of basic methods of gel electrophoresis and staining of DNA. Individuals should use this kit only in accordance with prudent laboratory safety precautions and under the supervision of a person familiar with such precautions. Use of this kit by unsupervised or improperly supervised individuals could result in injury.

Limited License: Polymerase chain reaction (PCR) is protected by patents owned by Hoffman-La Roche, Inc. The purchase price of this product includes a limited, non-transferable license under U.S. Patents 4,683,202; 4,683,195; and 4,965,188 or their foreign counterparts, owned by Hoffmann-La Roche Inc. and F. Hoffmann-La Roche Ltd. (Roche), to use only this amount of the product to practice the Polymerase Chain Reaction (PCR) and related processes described in said patents solely for the research, educational, and training activities of the purchaser when this product is used either manually or in conjunction with an authorized thermal cycler. No right to perform or offer commercial services of any kind using PCR, including without limitation reporting the results of purchaser's activities for a fee or other commercial consideration, is hereby granted by implication or estoppel. Further information on purchasing licenses to practice the PCR process may be obtained by contacting the Director of Licensing at The Perkin-Elmer Corporation, 850 Lincoln Center Drive, Foster City, California 94404 or at Roche Molecular Systems, Inc., 1145 Atlantic Avenue, Alameda, California 94501.

Printed material: The student instructions, pages 5-20, as well as the *Carolina*BLU™ staining protocol on page 25 may be photocopied as needed for use by your students.

REAGENTS, SUPPLIES, AND EQUIPMENT CHECKLIST

Included in all kits (21-1290, 21-1291, 21-1292, 21-1293, 21-1294, 21-1295):

Store in a -20°C freezer:

- ☐ 700 µl *CLF1*/*CLF2* primer/loading dye mix
- ☐ 700 µl *CLF1*/*Ds* primer/loading dye mix
- ☐ 130 µl tube pBR322/*Bst*NI markers (0.075 µg/µl)

Store in 4°C refrigerator:

- ☐ 2 tubes *Arabidopsis* seeds from self cross of *CLF*/*clf-2* heterozygotes

Store at room temperature:

- ☐ Growing tray
- ☐ Dome lid for growing tray
- ☐ Planting container with 6 cells
- ☐ 1 bag of potting soil, 4 cups
- ☐ 6-2 ml tubes Edward's buffer
- ☐ 6-2 ml tubes 100% Isopropanol
- ☐ 6-1 ml tubes Tris/EDTA (TE) buffer
- ☐ 25 *Ready-to-Go™ PCR Beads
- ☐ 5 ml mineral oil
- ☐ 15 plastic pestles
- ☐ Instructor's manual with reproducible student guide
- ☐ *CLF* CD-ROM

*Amplification and **Electrophoresis Kits (21-1292, 21-1293, 21-1294, 21-1295) also contain:*

- ☐ 2 vials of agarose, 2 g
- ☐ 1 bottle of 20x TBE, 150 ml
- ☐ 6 staining trays
- ☐ 1 bottle of 1 µg/ml ethidium bromide with MSDS, 250 ml (in kits with ethidium bromide)
- ☐ 12 latex gloves
- ☐ 1 bottle of *Carolina*BLU™ Gel & Buffer Stain, 7 ml (in kits with *Carolina*BLU™)
- ☐ 1 bottle of *Carolina*BLU™ Final Stain, 250 ml (in kits with *Carolina*BLU™)

Needed but not supplied:

- ☐ Micropipets and tips (1 µl to 1000 µl)
- ☐ Microcentrifuge tubes (minimum of 36)
- ☐ Microcentrifuge tube racks
- ☐ Microcentrifuge for 1.5-ml tubes
- ☐ Thermal cycler
- ☐ Water bath or heating block
- ☐ Electrophoresis chambers, trays, combs
- ☐ Electrophoresis power supplies
- ☐ UV transilluminator (to visualize ethidium bromide stained DNA)
- ☐ Permanent markers
- ☐ Containers with cracked or crushed ice
- ☐ Cool white fluorescent light (optional)
- ☐ Vortexer (optional)
- ☐ White light box (to visualize *Carolina*BLU™ stained DNA, optional)
- ☐ Camera or photo-documentary system (optional)

*Ready-to-Go™ PCR Beads incorporate *Taq* polymerase, dNTPs, and $MgCl_2$. Each bead is supplied in an individual 0.5–ml tube or a 0.2–mL tube.

**Electrophoresis reagents must be purchased separately for Kits 21-1290 and 21-1291.

Detecting a Transposon in *Arabidopsis*

CONTENTS

STUDENT LAB INSTRUCTIONS

INTRODUCTION

Throughout the first half of the 20th century, geneticists assumed that a stable genome was a prerequisite for faithfully transmitting genes from one generation to the next. Working at Cold Spring Harbor Laboratory in the post-WWII era, Barbara McClintock found quite a different story in maize (corn). She observed numerous "dissociations" – broken and ring-shaped chromosomes – and traced the source of these mutations to two related loci, "dissociator" (*Ds*) and "activator" (*Ac*).

Barbara McClintock at Cold Spring Harbor Laboratory, about 1950. Courtesy Cold Spring Harbor Laboratory Archives.

Equipped only with her maize crosses and a light microscope, she showed that *Ac* and *Ds* are mobile genetic elements that transpose, or jump. McClintock also offered genetic proof that *Ac* moves independently, but *Ds* depends on *Ac* for transposition. By showing that transposons may inactivate gene expression by inserting into a gene, or may reactivate expression by jumping out, McClintock explained color variegations, such as speckled kernels, that had intrigued botanists for centuries.

Today the *Ac*/*Ds* system is an important tool in gene discovery, allowing scientists to characterize genes for which no biological role is known. In a process known as transposon mutagenesis, *Ac* and *Ds* elements are crossed into a corn strain to produce *Ds* insertions in genes. The *Ac*/*Ds* mutagenesis system works well in a number of plants – including tobacco, tomato, and the model plant *Arabidopsis thaliana*.

This laboratory investigates a *curly leaf* (*clf*) mutant of *Arabidopsis thaliana* to analyze the molecular relationship between genotype and phenotype. *CLF* is involved in homeotic gene regulation, which controls the correct spatial and tissue-specific expression of genes during development. The recessive *clf-2* mutation seen in this lab was created through *Ds*

Wild-type (left) and *clf-2* mutant (right) *Arabidopsis* plants.

transposon mutagenesis and produces a dwarf phenotype with curly leaves, early flowers, and fused flower parts.

DNA is isolated from *Arabidopsis* plants and the *CLF* locus is amplified using polymerase chain reaction (PCR). Two primer sets, one specific for the wild-type allele, and one for the mutant allele bearing the *Ds* insertion, amplify DNA fragments that can readily be differentiated by agarose gel electrophoresis. Bioinformatics analysis unveils the structure of the *CLF* gene and the exact location of the *Ds* insertion in the gene.

Edwards, K., Johnstone, C., and Thompson, C. (1991). A Simple and Rapid Method for the Preparation of Plant Genomic DNA for PCR Analysis. *Nucleic Acids. Res.* **19**: 1349.

Fedoroff, N., Wessler, S., and McClure, M. (1983). Isolation of the Transposable Maize Controlling Elements *Ac* and *Ds*. *Cell* **35**: 235–242.

Feldman, K.A., Marks, M.D., Christianson, M.L., and Quantrano, R.S. (1989). A Dwarf Mutant of *Arabidopsis* Generated by T-DNA Insertion Mutagenesis. *Science* **243**: 1351-1354.

Goodrich, J., Puangsomlee, P., Martin, M., Long, D., Meyerowitz, E.M., and Coupland, G. (1997). A Polycomb-Group Gene Regulates Homeotic Gene Expression in *Arabidopsis*. *Nature* **386**: 44–51.

Hsieh, T-F., Hakim, O., Ohad, N., and Fischer, R.L. (2003). From Flour to Flower: How Polycomb-Group Proteins Influence Multiple Aspects of Plant Development. *Trends Plant Sci.* **8**: 439–445.

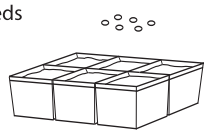
Kim, G-T., Tsukaya, H., and Uchimiya, H. (1998). The CURLY LEAF Gene Controls Both Division and Elongation of Cells During the Expansion of the Leaf Blade in *Arabidopsis thaliana*. *Planta* **206**: 175–183.

McClintock, B. (1951). Chromosome Organization and Genic Expression. *Cold Spring Harbor Symp. Quant. Biol.* **16**: 13–47.

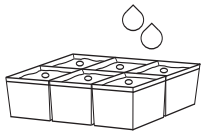
LAB FLOW

I. PLANT *ARABIDOPSIS* SEED

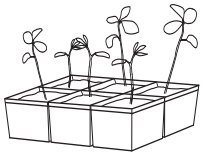
PLANT seeds



WATER seeds



3-4 WEEKS



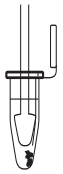
II. ISOLATE DNA FROM *ARABIDOPSIS*

ADD leaf tissue

GRIND

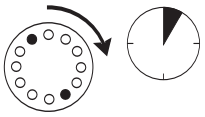
ADD Edwards' buffer

GRIND

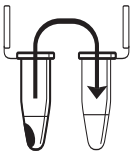


VORTEX

CENTRIFUGE

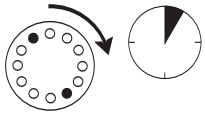


TRANSFER supernatant

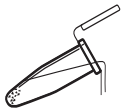


ADD and MIX Isopropanol

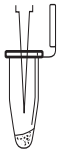
CENTRIFUGE



POUR OFF supernatant



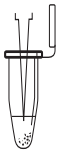
REMOVE supernatant



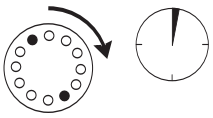
DRY

ADD TE buffer

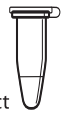
RESUSPEND DNA



CENTRIFUGE



leaf DNA extract



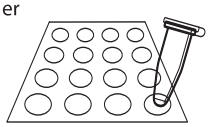
III. AMPLIFY DNA BY PCR

ADD primer/ loading dye mix

ADD DNA

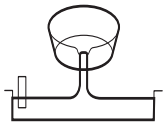
ADD mineral oil (if necessary)

AMPLIFY in thermal cycler



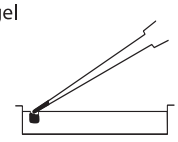
IV. ANALYZE PCR PRODUCTS BY GEL ELECTROPHORESIS

POUR gel

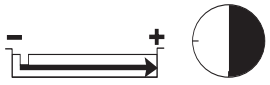


SET

LOAD gel



ELECTROPHORESE 130 volts

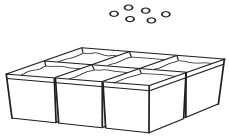


METHODS

I. PLANT *ARABIDOPSIS* SEED

To extract DNA from plant tissue, you must plant the Arabidopsis *seeds 3-4 weeks prior to DNA isolation and PCR. Depending upon growing conditions, you may observe the phenotypic differences between mutant and wild-type plants in as little as 2 weeks. For further information on cultivation, refer to The* Arabidopsis *Information Resource (TAIR) at http://www.arabidopsis.org.*

Reagents (at each student station)	**Supplies and Equipment**
Arabidopsis seeds from self cross of a *CLF/clf-2* heterozygote	Planting pots and tray Plastic dome lid or plastic wrap Potting soil Water



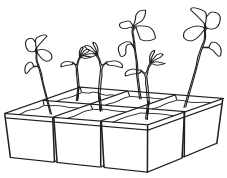
Arabidopsis seeds are very tiny and difficult to handle, so planting is not as simple as it may seem.

1. Moisten the potting soil. Fill planting pots evenly with soil, but do not pack tightly.
2. Fit pots into the tray, but leave one corner space empty to facilitate watering.
3. Carefully scatter seeds evenly on top of the soil.
   a. Fold a 4 × 4 inch sheet of paper in half.
   b. Place the seeds into the fold of the paper, and gently tap them onto the soil. Plant the tubes of seeds one at a time.
   c. Provide space between seeds, so they will grow better and plant phenotypes can be readily observed.
4. Cover pots with plastic dome lids or plastic wrap to maintain humidity during germination. (Remove covers 3–7 days after planting.)

To prevent the soil from drying out, keep a small amount of water in the tray at all times.

5. Add 1/2 inch of water to tray, using the empty corner space. Water regularly to keep soil damp, but do not allow soil to remain soggy.
6. Place planted seeds in refrigrator for 2-4 days. This "vernalization" is needed for optimum germination.

With 24-hour fluorescent lighting, phenotypes can be discerned in 2-3 weeks.



7. Grow the plants close to a sunny window at room temperature (20–22°C). For optimum growth, provide a constant (24 hours/day) fluorescent light source about 1 foot directly above the plants.
8. Harvest plant tissue for PCR when the wild-type and *clf-2* phenotypes become evident. You may wish to continue to grow the plants after you have harvested tissue for DNA isolation and amplification. The phenotypic differences between plants become more obvious over time, with the small *clf-2* mutants flowering well before the wild-type plants.

II. ISOLATE DNA FROM *ARABIDOPSIS*

Reagents (per 2 student stations)	**Supplies and Equipment**
2 *Arabidopsis* plants Edward's buffer, 2 ml Isopropanol, 2 ml Tris/EDTA (TE) buffer, 2 ml	2 plastic pestles Permanent marker 4 1.5 ml microcentrifuge tubes Micropipet and tips (100-1000 µl) Microcentrifuge tube racks Microcentrifuge Vortexer (optional) Container with cracked or crushed ice

Your instructor will assign you an *Arabidopsis* plant.

1. Obtain an *Arabidopsis* plant, and record its phenotype.
2. Cut two pieces of plant tissue approximately 1/4 inch in diameter. Place the tissue in a clean 1.5 ml tube, and label with the phenotype and your group number.
   a. While extracting DNA from one leaf may suffice, take multiple leaves from a small plant for an equivalent amount of leaf tissue.
   b. The *clf-2* mutant may be so small that you need to use the entire plant. If so, carefully remove all soil from the roots.

When fully ground, the sample should be a green liquid.

3. Twist a clean pestle against the inner surface of the 1.5 ml tube to *forcefully* grind the plant tissue for 1 minute.

Detergent in the Edward's buffer dissolves lipids of the cell membrane.

4. Add 400 µl of Edward's buffer to the tube. Grind briefly to remove tissue from the pestle and to liquify any remaining pieces of tissue.
5. Vortex the tube for 5 seconds, by hand or machine.
6. Place the tube in a balanced configuration in a microcentrifuge, and spin for 2 minutes to pellet any remaining cell debris. Spin longer if there is still unpelleted debris.
7. Transfer 350 µl of supernatant to a fresh tube. Maintain label for plant phenotype and group number. Be careful not to disturb the pelleted debris when transferring the supernatant. Discard old tube containing the precipitate.

This step precipitates nucleic acids, including DNA.

8. Add 400 µl of isopropanol to the tube of supernatant.
9. Mix by inverting the tube several times, and leave at room temperature for 3 minutes.

Centrifugation pellets the nucleic acids. The pellet may appear as a tiny teardrop-shaped smear or particles on the bottom side of the tube underneath the hinge. Don't be concerned if you can't see a pellet. A large or greenish pellet is cellular debris carried over from the first centrifugation.

10. Place your tube and those of other teams in a balanced configuration in a microcentrifuge, with the hinge of the cap pointed outwards. Spin for 5 minutes to pellet the DNA.
11. Carefully pour off and discard the supernatant from the tube. Then *completely* remove the remaining liquid with a medium pipet set at 100 µl.
12. Air dry the pellets by letting the tubes sit with caps open for 10 minutes. The remaining isopropanol will evaporate.
13. Add 100 µl of TE buffer to each tube. Dissolve DNA containing the

You will use 2.5 µl of the DNA extract for the PCR reactions in Part III. The crude DNA extract contains nucleases that will eventually fragment the DNA at room temperature. Keeping the sample cold limits this activity.

nucleic acid pellet by pipetting in and out. Take care to wash down the side of the tube underneath the hinge, where the pellet formed during centrifugation.

14. DNA may be used immediately or stored at –20°C until you are ready to continue with Part III. Keep the DNA on ice during use.

III. AMPLIFY DNA BY PCR

Reagents (at each student station)	**Supplies and Equipment**
Arabidopsis* DNA (from Part II) **CLF1/CLF2* primer/loading dye mix, 30 µl **CLF1/Ds* primer/loading dye mix, 30 µl Ready-To-Go™ PCR Beads **Shared Reagent Mineral oil, 5 ml (depending on thermal cycler) *Store on ice	Permanent marker Micropipets and tips (1-100 µl) Microcentrifuge tube rack Container with cracked or crushed ice Thermal cycler

1. Set up *CLF1/CLF2* reaction:
   a. Obtain a PCR tube containing a Ready-To-Go™ PCR Bead. Label with the primer set (*e.g.* "*CLF*") and your group number.
   b. Use a micropipet with a fresh tip to add 22.5 µl of the *CLF1/CLF2* primer/loading dye mix to your tube. Allow bead to dissolve.
   c. Use a micropipet with a fresh tip to add 2.5 µl of *Arabidopsis* DNA (from Part II). Insert the pipet tip directly into the PCR mix, eject the DNA, and insure that the tip comes out empty.

The primer loading dye mix will turn purple as the Ready-To-Go™ PCR Bead dissolves.

If the reagents become splattered on the wall of the tube, pool them by pulsing in a microcentrifuge or by sharply tapping the tube bottom on the lab bench.

2. Set up *CLF1/Ds* reaction:
   a. Obtain a PCR tube containing a Ready-To-Go™ PCR Bead. Label with the primer set (e.g. "*Ds*") and your group number.
   b. Use a micropipet with a fresh tip to add 22.5 µl of the *CLF1/Ds* primer/loading dye mix to your tube. Allow bead to dissolve.
   c. Use a micropipet with a fresh tip to add 2.5 µl of *Arabidposis* DNA (from Part II). Add DNA directly into the PCR mix.

Mineral oil prevents the PCR mix from evaporating and condensing on the tube cap during cycling. Most modern thermal cyclers have heated lids that prevent condensing and DO NOT require the addition of mineral oil.

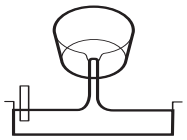
3. If necessary, add one drop of mineral oil to the top of the reactants in the PCR tubes. Be careful not to touch the dropper tip to the tube or reactants, or subsequent reactions will be contaminated with DNA from your preparation.
4. Store samples on ice until you are ready to begin thermal cycling.
5. Program the thermal cycler for 30 cycles of the following profile. The program may be linked to a 4°C hold step after the 30 cycles are completed.

Denaturing step:	94°C	30 seconds
Annealing step:	65°C	30 seconds
Extending step:	72°C	30 seconds

6. After cycling, store the amplified DNA at –20°C.

IV. ANALYZE PCR PRODUCTS BY GEL ELECTROPHORESIS

Reagents (per 2 student stations)	**Supplies and Equipment**
Arabidopsis* PCR products (from Part III) 1× TBE, 300 ml 2% agarose in 1× TBE, 50 ml **Shared Reagents *pBR322/*Bst*NI marker, 130 µl Ethidium bromide (1 µg/ml), 250 ml or *Carolina*BLU™ Gel & Buffer Stain, 7 ml *Carolina*BLU™ Final Stain, 250 ml *Store on ice	Micropipets and tips (1-20 µl, 20-100 µl) 1.5 ml microcentrifuge tube rack 1 1.5 ml microcentrifuge tube Gel electrophoresis chamber Power supply Staining trays Latex gloves UV transilluminator for use with ethidium bromide) White light transilluminator (for use with *Carolina*BLU™) Digital or instant camera (optional) Water bath (60°C) Container with cracked or crushed ice

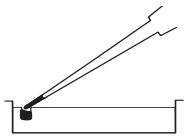


Avoid pouring an overly thick gel, which is more difficult to visualize. The gel will become cloudy as it solidifies.

1. Seal the ends of the gel-casting tray with masking tape, and insert a well-forming comb.
2. Pour 2% agarose solution to a depth that covers about 1/3 the height of the open teeth of the comb.
3. Allow the gel to solidify completely. This takes approximately 20 minutes.
4. Place the gel into the electrophoresis chamber, and add enough 1× TBE buffer to cover the surface of the gel.

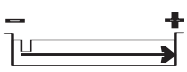
Do not add more buffer than necessary. Too much buffer above the gel channels electrical current over the gel, increasing running time.

5. Carefully remove the combs, and add additional 1× TBE buffer to just cover and fill in wells – creating a smooth buffer surface.
6. Combine both PCR products (*CLF1/CLF2* and *CLF1/Ds*) together in a fresh 1.5 ml tube. Pipet in and out to mix. (If you used mineral oil during PCR, pierce your pipet tip through the layer of mineral oil to withdraw the PCR products and leave the mineral oil behind in the original tube.)



Expel any air from the tip before loading. Be careful not to push the tip of the pipet through the bottom of the sample well.

7. Use a micropipet with a fresh tip to add 30 µl of the combined PCR sample/loading dye mixture into your assigned well of a 2% agarose gel.
8. Load 20 µl of the molecular weight marker (pBR322/*Bst*NI) into one well. (Alternatively, use appropriate amount of a 100-bp DNA ladder.)



9. Run the gels at 130 V for approximately 30 minutes. Adequate separation will have occurred when the cresol red dye front has moved at least 50 mm from the wells.

10. Stain the gel in ethidium bromide or *Carolina*BLU™:

    a. For ethidium bromide, stain 10-15 minutes. Decant stain back into storage container for reuse, and rinse gel in tap water. Use gloves when handling ethidium bromide solution and stained gel, or anything that has ethidium bromide on it. Ethidium bromide is a known mutagen and care should be taken when using and disposing of it.

    b. For *Carolina*BLU™ staining, follow directions in the Instructor Planning section.

Destaining the gel for 5-10 minutes in tap water leeches unbound ethidium bromide from the gel, decreasing background and increasing contrast of the stained DNA.

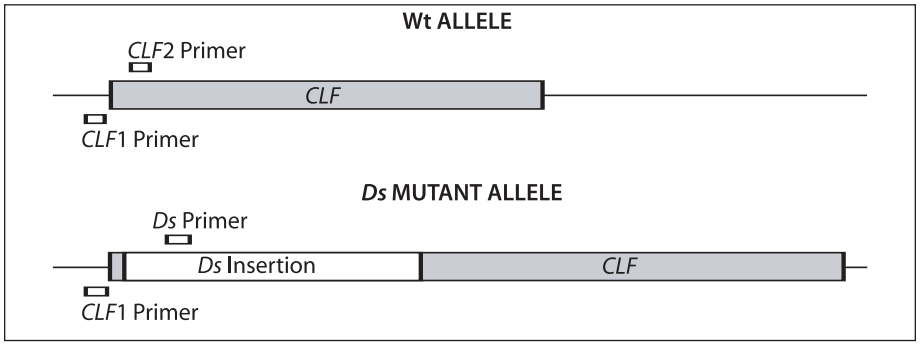
11. View gel using transillumination, and photograph using a digital or instant camera. Draw an image of the bands on the gel if a camera is not available.

Transillumination, where the light source is below the gel, increases brightness and contrast.

RESULTS & DISCUSSION

This laboratory investigates the *CURLY LEAF (CLF)* gene of *Arabidopsis thaliana* to show the molecular relationship between genotype and phenotype. The *clf-2* mutation seen in this lab was created by *Ds* transposon mutagenesis and produces a dwarf phenotype with curly leaves, early flowers, and fused flower parts.

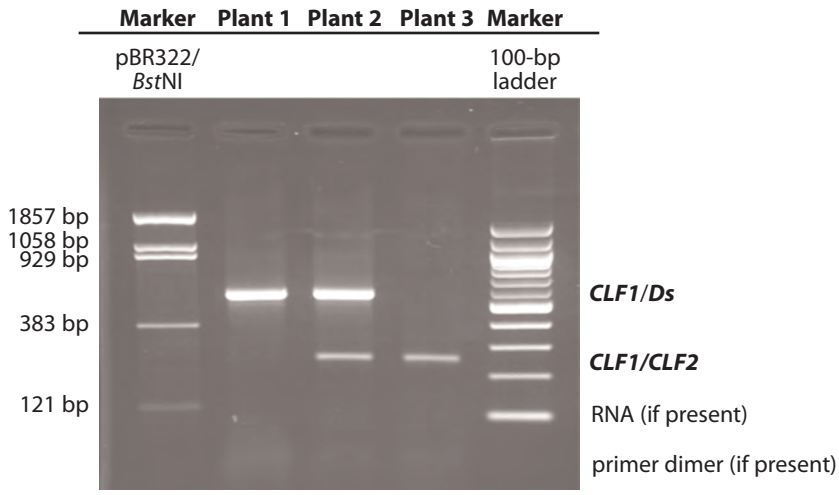
DNA was isolated from *Arabidopsis* plants, and the *CLF* locus was amplified using polymerase chain reaction (PCR). Since the *Ds* insertion at the *CLF* locus is too large to amplify across, two sets of primers were used. One set (*CLF1*/*CLF2*) straddles the *Ds* insertion site and amplifies only the wild-type allele. The second set (*CLF1*/*Ds*) amplifies only the insertion allele, with one primer located in the *CLF* gene and one located in the *Ds* transposon.



PCR amplification of the *CLF* locus.

1. Describe the purpose of each of the following steps or reagents used in DNA isolation (Part II):
   a. grinding with pestle.
   b. Edward's buffer.
   c. Tris-EDTA (TE) buffer.
2. What is the purpose of performing each of the following PCR reactions:
   a. *CLF1/CLF2*?
   b. *CLF1/Ds*?
3. The photo on the following page shows an example of expected results. Place Plants 1-3 of the example gel in the proper cells of the chart below, and complete the entries for genotype and phenotype. (Remember that your own samples can yield any of the combinations for *CLF1/CLF2* and *CLF1/Ds* in the table below.)

246 bp (*CLF1*/*CLF2*)	539 bp (*CLF1*/*Ds*)	Plant Genotype	Plant Phenotype	Sample
present	present			
present	absent			
absent	present			



4. View the image of the stained gel containing your sample and those of other students. Orient the photograph with wells at the top. Interpret each lane of the gel.

   a. Scan across the photograph of your gel to get an impression of what you see in each lane. You should notice that virtually all experiment lanes contain one or two prominent bands.

   b. Now locate the lane containing the pBR322/*Bst*NI marker on the left hand side of the gel. Working from the well, locate the bands corresponding to each restriction fragment: 1,857 bp, 1,058 bp, 929 bp, 383 bp, and 121 bp (may be faint or not visible). (If you used the 100-bp marker, locate the lane containing the 100-bp ladder, with the fastest migrating band equal to 100 bp and each successive band 100 bp larger.)

   c. The amplification product of the mutant *clf-2* allele (539 bp) should align between the 383-bp and 929-bp fragments of the pBR322/*Bst*NI marker (or between the 500-bp and 600-bp fragments of the 100-bp ladder).

   d. The amplification product of the wildtype *CLF* allele (246 bp) should align between the 121-bp and 383-bp fragments of the pBR322/*Bst*NI marker (or between the 200-bp and 300-bp fragments of the 100-bp ladder).

   e. It is common to see one or two diffuse (fuzzy) bands of RNA and/or primer dimer at the bottom of the gel. RNA may be found at approximately the position of the 121-bp fragment of the pBR322/*Bst*NI marker (or the 100-bp marker of the 100-bp ladder). RNA is the largest component of nucleic acid isolated from plant tissue. Primer dimer is an artifact of the PCR reaction that results from two primers overlapping one another and amplifying themselves. Primer dimer is approximately 50 bp, and should be in a position ahead of the 121-bp marker fragment.

   f. Additional faint bands, at other positions in the gel, occur when the primers bind to chromosomal loci other than the *CLF* locus, giving rise to "nonspecific" amplification products.

5. Based on analysis of the agarose gel, what is the genotype of your plant?
6. How would you interpret a lane in which you observe primer dimer, but not the *clf-2* or wild-type alleles?
7. Would you classify the *clf-2* mutation as recessive or dominant? Explain your reasoning.
8. Based on the genotypic distribution of the plants analyzed by the class, what genotype was the parental plant? Why may the genotype distribution observed by your class deviate from what is expected under Mendel's laws of inheritance?

BIOINFORMATICS

Biological information is encoded in the nucleotide sequence of DNA. Bioinformatics is the field that identifies biological information in DNA using computer-based tools. Some bioinformatics algorithms aid the identification of genes, promoters, and other functional elements of DNA. Other algorithms help determine the evolutionary relationships between DNA sequences.

Because of the large number of tools and DNA sequences available on the Internet, experiments done *in silico* ("in silicon," or computer) now complement experiments done *in vitro* (in glass, or test tube). This movement between biochemistry and computation is a key feature of modern biological research.

In Part I you will use the Basic Local Alignment Search Tool (BLAST) to identify sequences in biological databases and to make predictions about the outcome of your experiments. In Part II you will identify the sequences of the *CLF* protein and the *CLF1/CLF2* amplicon. In Part III you will discover the chromosome location of the *CLF* gene. In Part IV you will determine where the *Ds* transposon is inserted into the *CLF* gene to produce the *clf-2* mutation. In Part V you will learn about the function and the evolutionary history of the *CLF* protein.

NOTE: The links in these bioinformatics execises were correct at the time of printing. However, links and labels within the NCBI Internet site change occasionally. When this occurs, you can find updated exercises at http://bioinformatics.dnalc.org.

I. Use *BLAST* to Find DNA Sequences in Databases (Electronic PCR)

1. Initiate a BLAST search.
   a. Open the Internet site of the National Center for Biotechnology Information (NCBI), www.ncbi.nlm.nih.gov/.
   b. Click on *BLAST* in the top speed bar.
   c. Click onthe link to *nucleotide BLAST* under the heading *Basic BLAST*.

d. Enter the *CLF1/CLF2* primer set sequences into the *Search* window. These are the query sequences.

The following primer sets were used in the experiment:	
CLF1	5'-TTAACCCGGACCCGCATTTGTTTCGG-3' (Forward Primer)
CLF2	5'-AGAGAAGCTCAAACAAGCCATCGA-3' (Reverse Primer)
CLF1	5'-TTAACCCGGACCCGCATTTGTTTCGG-3' (Forward Primer)
Ds	5'-GTCGGCGTGCGGCTGGCGGCG-3' (Reverse Primer)

e. Omit any non-nucleotide characters from the window, because they will not be recognized by the BLAST algorithm.

f. Under Choose *Search Set*, select the *Nucleotide collection(nr/nt)* database from the drop-down menu.

g. Under *Program Selection*, optimize for somewhat similar sequences by selecting *blastn*.

h. Click on *BLAST!* and the query sequences are sent to a server at the National Center for Biotechnology Information in Bethesda, Maryland. There, the BLAST algorithm will attempt to match the primer sequences to the millions of DNA sequences stored in its database. While searching, a page showing the status of your search will be displayed until your results are available. This may take only a few seconds, or more than a minute if a lot of other searches are queued at the server.

2. The results of the BLAST search are displayed in three ways as you scroll down the page:

   a. First, a graphical overview illustrates the significance of matches and how they align with the query sequence. Matches of differing lengths are coded by color.

   b. This is followed by a list of *significant alignments*, or hits, with links to *Accession* information.

   c. Next, is a detailed view of each primer sequence (*query*) aligned to the nucleotide sequence of the search hit (*subject*). Notice that the first match to the forward primer (nucleotides 1-26), and to the reverse primer (nucleotides 27-50) are within the same *Accession*. Which positions do the primers match in the subject sequence?

   d. What can you say about the rest of the alignments in the list?

3. Calculate the predicted length of the product that the primer set will amplify in a PCR reaction (*in vitro*) as follows:

   a. In the list of *significant alignments*, notice the scores in the *E-Value* column on the right. The *Expectation* or *E-Value* is the number of alignments with the query sequence that would be expected to occur by chance in the database. The lower the *E-Value* the higher the probability that the hit is related to the query. What does the *E-value* of *2e-4* mean?

b. Note the names of any *significant alignments* that have *E-values* less than 0.1. Do they make sense?

c. Scroll down to the *Alignments* section to see exactly where the two primers have landed in a subject sequence.

d. The lowest and highest nucleotide positions in the subject sequence indicate the borders of the amplified sequence. Subtracting one from the other gives the difference between the two coordinates.

e. However, the actual length of the fragment *includes* both ends, so add 1 nucleotide to the result to determine the exact length of the PCR product amplified by the two primers.

II. Identify the *CLF* Amino Acid Sequence and the *CLF1/CLF2* Amplicon

1. Click on the *Accession* link to open the sequence datasheet for this hit.
2. At the top of the report is basic information about the sequence, including its basepair length, database accession number, source, and references. Describe the source and size of the sequence in which your BLAST hit is located.
3. The middle part contains annotations of sequence features, with their beginning and ending nucleotide positions (xx .. xx). Identify the feature(s) contained between the nucleotide positions that match the primers, as determined in I.3.d above. What do they tell you about the name of the locus?
4. Look at the CDS (coding sequence) entry. "Join" shows the coordinates of coding exons that are spliced together for translation into a polypeptide chain. Each entry, XX..XX, gives the first and last nucleotides of one exon. How many exons are in the *CLF* gene? Where does the reverse primer end?
5. The final part of the CDS entry is "translation" which lists abbreviations for the amino acids specified by the coding sequence. Copy the entire amino acid sequence, and paste it into a text document for future use.
6. At the bottom of the report is the entire nucleotide sequence of the Chromosome 2 clone. Highlight all the nucleotides from the beginning of the forward primer to the end of reverse primer. Copy this sequence and paste it into a text document for future use. Delete all non-nucleotide characters. What is this sequence?

III. Use *Map Viewer* to Determine the Chromosome Location of the *CLF* Gene

1. Return to the NCBI home page, then click on *Map Viewer* located in the *Hot Spots* column on the right.

2. Scroll down the list in the table to the right and click on the "triangle" icon next to flowering plants, then *Eudicots*. Next, click on the "B" icon under the *Tools* header for the *Arabidopsis thaliana*. If more than one build is displayed, select the one with the highest number, as this will be the most recent version.
3. The primers are too short to produce a result in the *Map Viewer*. So copy the sequence of the 246-basepair amplicon from your text document, and paste it into the search window.
4. Click on *Begin Search* to send the sequence to the NCBI server.
5. Click on *Format* to retrieve the results.
6. Click on *[Thale cress genome view*] in the list of *Other reports* at the top of the page to see the chromosome location of the BLAST hit. On what chromosome have you landed?
7. Click on the marked chromosome to move to the *CLF* locus.
8. Click on the small blue arrow labeled *Genes seq* to display genes. The *CLF* gene occupies most of the field of the close-up view of *Arabidopsis* chromosome 2. Move the *zoom out* toggle to 1/10,000 and 1/1,000 to see more of the *CLF* gene.
9. For each gene diagram, exons (thick bars) are connected by introns (thin lines). Where does the red amplicon match the *CLF* gene?
10. Click the names of several neighboring genes, and follow links for more information about them. Can you figure out what the prefix At2g means?

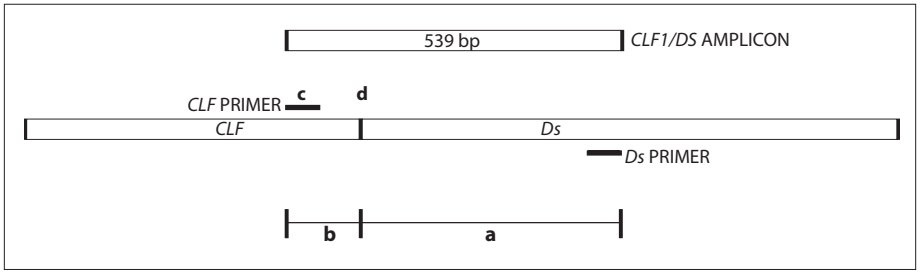
IV. Determine the Insertion Site of the *Ds* Transposon

1. The photograph of a gel of the PCR results shows that the *CLF1/Ds* amplicon, from *clf-2* mutants, is approximately 540 nucleotides long. This amplicon is composed partly of sequence from the *CLF* gene and partly of sequence from the *Ds* transposon. After determining how many nucleotides of the amplicon are from the *Ds* transposon, you can estimate the position at which *Ds* inserted into the *CLF* gene.
2. Repeat a *nucleotide-nucleotide BLAST* search with the *CLF1/Ds* primers. What hits match the *CLF* primer? Do these hits also have matches to the *Ds* primer? What can you conclude?
3. Now examine the hits for the *Ds* primer. Do they make sense? What do they have in common?
4. Label the following diagram with your answers to Parts a-c below to help you determine the *Ds* insertion site in the *CLF* gene (d.).
   a. Scroll to the *Alignments* section and focus on one of the complete *Ds2* sequences. How far into the *Ds2* sequence is the *Ds* primer?
   b. How many *CLF* nucleotides are in the *CLF/Ds* amplicon?

c. What is the first nucleotide position of the *CLF1* primer in the *CLF* amplicon?

d. What is the insertion position of *Ds*?

5. Return to the mRNA or CDS entries on the datasheet for the *CLF* gene, to determine into which exon or intron the *Ds* transposon is inserted.

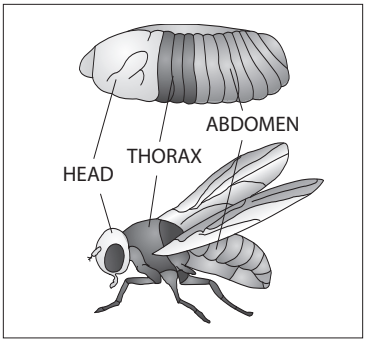


V. Use BLAST to Determine the Function of the *CLF* Protein

1. Return to the BLAST page at NCBI, www.ncbi.nlm.nih.gov/BLAST/. This time, click on *protein BLAST*.
2. Copy the *CLF* amino acid sequence from your text file, and paste it into the *Search* window.
3. Click on *BLAST!* to send the amino acid sequence to the NCBI server.
4. An algorithm quickly scans the query sequence to identify functional domains (regions of the protein) that are conserved in different organisms. Follow links to get some quick information about a domain. What do you find?
5. Once the BLAST results appear, examine the results.
6. What can you tell from the graphical overview?
7. Scroll down to the list of *significant alignments*. What can you conclude from the *E-values*, the titles of the hits, and the *Accession* links?
8. Scroll down to the *Alignments* to see just how well the *CLF* amino acid sequence matches that from other organisms. Follow *Accession* links to sequence datasheets – then on to journal abstracts and articles – to gain insight into how *CLF* and other homeotic proteins control development throughout the plant and animal worlds.

Homeotic Genes in Development

Homeotic genes were first discovered in *Drosophila* but have since been found in vertebrates and plants. Homeotic genes are required for the correct spatial and tissue-specific expression of genes that control early development of an organism. During development in *Drosophila*, the embryo is divided into smaller and smaller domains through the action of a hierarchy of homeotic genes. The first genes in this hierarchy establish the anterior/posterior and dorsal/ventral axes of the embryo. The next



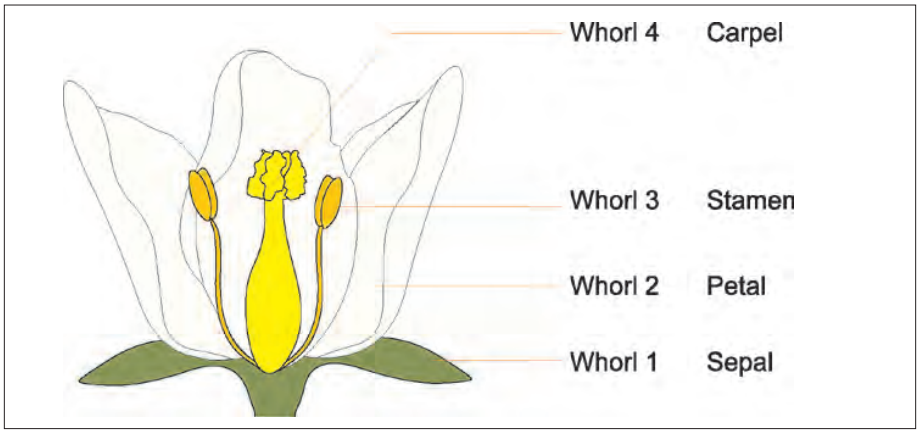

Antennapedia

three series of genes – the gap, pair rule, and segment polarity genes – divide the anterior/posterior axis into segments. The last genes are expressed throughout the remainder of development and adult life.

In *Drosophila* the combined activities of the homeotic proteins specify the identities of the three body segments: head, thorax, and abdomen. In homeotic gene mutants, one body part is exchanged for a different one. (For example, the *Antennapedia* mutant has legs in the place of antennae.)

In order for an organism to develop normally, correct patterns of homeotic gene expression must be maintained through multiple mitotic divisions that occur during development. These patterns can be thought of as a molecular memory system. Genes related to *Antennapedia* and *Bithorax* establish initial patterns of homeotic gene expression. These patterns are maintained later in development by genes of the Polycomb-Group (*Pc-G*) and trithorax-Group (*trx-G*).

Arabidopsis flowers contain four organs – sepals, petals, stamens, and carpels – that are organized into concentric whorls (below). Similar to *Drosophila* segments, the identities of these whorls are established by the combined action of homeotic genes.



Arrangement of flower organs.

The *Arabidopsis* homeotic gene *AGAMOUS* is normally expressed in whorls 3 and 4 of the flower, and is required for correct development of stamens and carpels. It is not expressed in whorls 1 and 2 of the flower, which develop into sepals and petals. This is because the *CLF* protein represses *AGAMOUS* expression in these whorls. Thus, like other Polycomb-Group proteins, *CLF* helps maintain patterns of homeotic gene expression after whorl identity has already been established. However, plants that are homozygous for the *clf-2* mutation fail to repress *AGAMOUS* and exhibit partial transformation of sepals and petals into stamens and carpels. Although the exact mechanism is not known, the ectopic (unusual) expression of *AGAMOUS* is also responsible for the more visible elements of the *clf-2* phenotype – curled leaves, dwarf size, and early flowering.

INFORMATION FOR INSTRUCTOR

CONCEPTS AND METHODS

This laboratory can help students understand several important concepts of modern biology:

- The relationship between genotype and phenotype.
- The integration of *in vitro* experimentation and *in silico* computation.
- The use of transposable elements to mutagenize and tag genes.
- The role of homeotic genes in plant and animal development.

The laboratory uses several methods from modern biological research:

- DNA extraction and purification.
- Polymerase chain reaction (PCR).
- Gel electrophoresis.
- Bioinformatics.

INSTRUCTOR PLANNING, PREPARATION, AND LAB FINE POINTS

The following table will help you to plan and integrate the four parts of the experiment.

Part		Day	Time	Activity	
I.	Plant A*rabidopsis* seeds	3-4 weeks prior to lab	15-30 min.	**Lab:**	Plant *Arabidopsis* seeds
II.	Isolate DNA	1	30 min. 30-60 min.	**Pre-lab:** **Lab:**	Set up student stations Isolate *Arabidopsis* DNA
III.	PCR-amplify DNA	2	30-60 min. 15-30 min. 70+ min.	**Pre-lab:** **Lab:** **Post-lab:**	Set up student stations Set up PCR reactions Amplify DNA in thermal cycler
IV.	Analyze Amplified DNA	3 4	30 min. 30 min. 30+ min. 20+ min. 20 min. to overnight 20 min.	**Lab:** **Post-lab:**	Prepare agarose solution; cast gels Load DNA samples into gels Electrophorese samples Stain gels De-stain gels Photograph gels

I. PLANT *ARABIDOPSIS* SEED

Arabidopsis *seeds must be planted 3-4 weeks before the date anticipated for DNA extraction and amplification by PCR.* Two 1/4 inch diameter leaf disks are required for each experiment, but multiple small leaves and even whole plantlets can be used.

Light fixtures for growing *Arabidopsis* should be fitted with at least two 40-watt cool white fluorescent bulbs.

II. ISOLATE DNA FROM *ARABIDOPSIS*

The reagents for DNA isolation are pre-aliquoted into six portions, each of which is shared between two student teams.

Assign each team a number at the outset of the experiment. This will make it easier to mark and identify the several types of small tubes used in the experiment

Make sure that each student team extracts DNA from either wild-type or *clf-2* mutant plants and that both types of plants are being examined by the class as a whole.

The cell walls of living plant tissue typically are broken up by grinding with a mortar and pestle. This can be accomplished directly in a 1.5 ml tube with the plastic pestle provided in the kit.

Pre-lab Set Up (per 2 student teams)

2 *Arabidopsis* plants
Edward's buffer, 2 ml
Isopropanol, 2 ml
Tris/EDTA (TE), 1 ml

4 1.5 ml microcentrifuge tubes
Permanent marker
2 plastic pestles
Micropipet and tips (100-1000 µl)
Microcentrifuge tube rack
Container with cracked or crushed ice

Shared Items

Microcentrifuge
Vortexer (optional)

III. AMPLIFY DNA BY PCR

The *Ds* insertion at the *CLF* locus is too large to amplify across, so a single primer set cannot amplify both the wild-type and insertion alleles. Thus, this experiment amplifies the wild-type and insertion alleles in separate PCR reactions using two different sets of primers (see RESULTS & DISCUSSION). One primer set (*CLF1/CLF2*) spans the *Ds* insertion site and thus amplifies only the wild-type *CLF* allele. The *CLF2* primer is moved out of amplifiable range in chromosomes with a *Ds* insertion. So, to amplify the insertion allele, the *CLF1* primer is paired with a second primer, *Ds*, which is located within the *Ds* transposon. Since the amplified fragments expected in the two PCR reactions differ sufficiently in size the results of the two PCR reactions are mixed prior to electrophoresis to simplify analysis of the three genotypes (*CLF/CLF*, *CLF/clf-2*, and *clf-2/clf-2*).

Each Ready-To-Go™ PCR Bead contains reagents so that when brought to a final volume of 25 µl the reaction contains 2.5 units of *Taq* DNA polymerase, 10 mM Tris-HCl (pH 9.0), 50 mM KCl, 1.5 mM $MgCl_2$, and 200 µM of each dNTP.

The lyophilized *Taq* DNA polymerase in the Ready-To-Go™ PCR beads become active immediately upon addition of the primer/loading dye mix. In the absence of thermal cycling, "nonspecific priming" at room temperature allows the polymerase to begin generating erroneous products, which can show up as extra bands in gel analysis. *Therefore, work quickly. Be sure the thermal cycler is set and have all experimenters set up their PCR reactions as a coordinated effort. Add primer/loading dye mix to all reaction tubes, then add each*

student template, and begin thermal cycling as quickly as possible. Hold reactions on ice until all are ready to load into the thermal cycler.

Each primer/loading dye mix incorporates the appropriate primer pair (0.26 picomoles/µl of each primer), 13.8% sucrose, and 0.0081% cresol red. The inclusion of loading dye components, sucrose and cresol red, allows the amplified product to be directly loaded into an agarose gel for electrophoresis. The primer/loading dye mix may collect in the tube caps during shipping; pool the reagent by spinning the tubes briefly in a microcentrifuge or by tapping the tube ends on the desktop.

PCR amplification from crude cell extracts is biochemically demanding, and requires the precision of automated thermal cycling. However, amplification of the *CLF* locus is not complicated by the presence of repeated units. Therefore, the recommended amplification times and temperatures will work adequately for all types of thermal cyclers.

Pre-lab Set Up (per student team)

Arabidopsis DNA, from Part II (store on ice)
30 µl *CLF1/CLF2* primer/loading dye mix (thaw and store on ice)
30 µl *CLF1/Ds* primer/loading dye mix (thaw and store on ice)
2 Ready-To-Go™ PCR Beads (in PCR tubes)

Permanent marker
Micropipet and tips (1-100 µl)
Microcentrifuge tube rack
Container with cracked or crushed ice

Shared Items

Thermal cycler
Mineral oil, 5 ml (depending on thermal cycler)

IV. ANALYZE AMPLIFIED DNA BY GEL ELECTROPHORESIS

Enough reagents are provided to prepare and electrophorese 6 gels.

Prepare a 1X concentration of TBE by adding 75 ml of 20X concentrated stock into 1,425 ml of deionized or distilled water. Mix thoroughly.

Prepare a 2% agarose solution by adding 2 g of agarose to 100 ml of 1X TBE in a 500 ml flask or beaker. Heat the flask or beaker in a microwave oven (approximately 4 minutes) or in a boiling water bath (approximately 15 minutes) until the agarose is completely dissolved. You should no longer see agarose particles floating in solution when the beaker is swirled. Allow the agarose to cool to approximately 60°C, and hold at this temperature in a hot water bath. Cover beaker or flask with aluminum foil, and skim any polymerized "skin" off the top of the solution before pouring. Notice there is enough agarose to pore 6-7, 50 ml gels. There is enough marker to run 6 gels.

The cresol red and sucrose in the primer mix function as loading dye, so that amplified samples can be loaded directly into an agarose gel. This is a nice time saver. However, since it has relatively little sugar and cresol red, this loading dye is more difficult to use than typical loading dyes. So, encourage students to load carefully.

Plasmid pBR322 digested with the restriction endonuclease *Bst*NI is an inexpensive marker and produces fragments that are useful as size markers in this experiment. The size of the DNA fragments in the marker are 1,857 bp, 1,058 bp, 929 bp, 383 bp, and 121 bp. Use 20 µl of a 0.075 µg/µl stock solution of this DNA ladder per gel. Other markers or a 100-bp ladder may be substituted.

View and photograph gels as soon as possible after appropriate staining/destaining. Over time, the small-sized PCR products will diffuse through the gel and lose sharpness. (If absolutely necessary, gels can be wrapped in plastic wrap and stored at 4°C for up to 24 hrs.)

Pre-lab Set Up (per 2 student teams)

Arabidopsis PCR products from Part III (store on ice)
2% agarose in 1X TBE (hold at 60°C), 50 ml per gel
1X TBE buffer, 300 ml per gel

1.5 ml microcentrifuge tubes
Micropipets and tips (1-20 µl; 20-100 µl)
Microcentrifuge tube rack
Electrophoresis chamber and power supply
Latex gloves
Staining tray
Container with cracked or crushed ice

Shared Items

Water bath for agarose solution (60°C)
pBR322/*Bst*NI markers (thaw and store on ice)
Ethidium bromide (1 µg/ml), 250 ml
or
*Carolina*BLU™ Gel & Buffer Stain, 7 ml and *Carolina*BLU™ Final Stain
Transilluminator with camera

*Carolina*BLU™ STAINING

POST-STAINING

1. Cover the electrophoresed gel with the *Carolina*BLU™ Final Stain and let sit for 20–30 minutes. Agitate gently (optional).
2. After staining, pour the stain back into the bottle for future use. (The stain can be used 6–8 times.)
3. Cover the gel with deionized or distilled water to destain. Chloride ions in tap water will cause the staining to fade.
4. Change the water 3 or 4 times over the course of 30–40 minutes. Agitate the gel occasionally.
5. Bands that are not immediately present will become more apparent with time and will reach their maximum visibility if the gel is left to destain overnight in just enough water to cover the gel. Gels left overnight in a large volume of water may destain too much.

PRE-STAINING

*Carolina*BLU™ can also be used to stain the DNA while it is being electrophoresed. Pre-staining will allow students to visualize their results prior to the end of the gel run. However, post-staining is still required for optimum viewing.

To pre-stain the gel during electrophoresis, add *Carolina*BLU™ Gel and Buffer Stain in the amounts indicated in the tables below. Note that the amount of stain added is dependent upon the voltage used for electrophoresis. *Do not use more stain than recommended. This may precipitate the DNA in the wells and create artifact bands.*

Gels containing *Carolina*BLU™ may be prepared one day ahead of the lab day, if necessary. However, gels stored longer tend to fade and lose their ability to stain DNA bands during electrophoresis.

Use the table below to add the appropriate volume of *Carolina*BLU™ stain to the agarose gel:

Voltage	Agarose Volume	Stain Volume
<50 Volts	30 ml	40 µl (1 drop)
	200 ml	240 µl (6 drops)
	400 ml	520 µl (13 drops)
>50 Volts	50 ml	80 µl (2 drops)
	300 ml	480 µl (12 drops)
	400 ml	640 µl (16 drops)

Use the table below to add the appropriate volume of *Carolina*BLU™ stain to 1× TBE buffer:

Voltage	Agarose Volume	Stain Volume
<50 Volts	500 ml	480 µl (12 drops)
	3000 ml	2.88 ml (72 drops)
>50 Volts	500 ml	960 µl (24 drops)
	2600 ml	5 ml (125 drops)

ANSWERS TO DISCUSSION QUESTIONS

1. Describe the purpose of each of the following steps or reagents used in DNA isolation (Part II):
   a. **Grinding with the pestle breaks the cell walls of the *Arabidopsis* tissue.**
   b. **The detergent component of Edward's buffer, sodium dodecyl sulfate, dissolves lipids that compose the cell membrane.**
   c. **Tris-EDTA (TE) buffer provides conditions for stable storage of DNA. Tris provides a constant pH of 8.0, while EDTA binds cations (positive ions) that are required for DNase activity.**
2. What is the purpose of performing each of the following PCR reactions:
   a. *CLF1/CLF2*? **Detects the presence of the wild-type *CLF* allele.**
   b. *CLF1/DS*? **Detects the presence of the mutant *clf-2* allele bearing a *Ds* transposon insertion.**
3. Place Plants 1-3 of the example gel in the proper cells of the chart below, and complete the entries for genotype and phenotye.

246 bp (*CLF1/CLF2*)	539 bp (*CLF1/Ds*)	Plant Genotype*	Plant Phenotype	Sample
present	present	**+/-**	**wild-type**	**Plant 2**
present	absent	**-/-**	**wild-type**	**Plant 3**
absent	present	**+/+**	**mutant**	**Plant 1**

*"+" indicates presence of transposon. "-" indicates its absence.

5. Based on analysis of the agarose gel, what is the genotype of your plant? **If you observe only the smaller band, the plant is homozygous for the wild-type allele (-/-). If you observe only the larger band, the plant is homozygous for the mutant *clf-2* allele (+/+). If you observe both bands, the plant is heterozygous (+/-).**
6. How would you interpret a lane in which you observe primer dimer, but not the *clf-2* or wild-type alleles? **The presence of primer dimer confirms that the reaction contained all the components necessary for amplification, but there was insufficient template to amplify target sequences.**
7. Would you classify the *clf-2* mutation as recessive or dominant? Explain your reasoning. **The *clf-2* mutation is recessive. Plants that are heterozygous do not demonstrate the mutant phenotype. Only plants that are homozygous for the *clf-2* mutation demonstrate the curly leaf phenotype.**
8. Based on the genotypic distribution of the plants analyzed by the class, what genotype was the parental plant? Why may the genotype distribution observed by your class deviate from what is expected under Mendel's laws of inheritance? **First, count the number of +/+, +/-, and -/- plants, and express it as a ratio. Then, draw Punnett squares for the six kinds of parental crosses – which produce three different geneotypic ratios. The observed ratio of +/+, +/-, and -/- plants should be closest to the 1:2:1 ratio predicted by a cross between two heterozygous parents. Small sample size or bias toward picking one phenotype will cause deviations from the expected ratio.**

   (*Arabidopsis thaliana* has the ability to self-pollinate. The seeds for this experiment, in fact, came from a parental plant that self-pollinated. Think about the kind of self-crosses that are possible for plants of each of the three genotypes!)

ANSWERS TO BIOINFORMATICS QUESTIONS

I.2. c. Which positions do the primers match in the subject sequence? **Positions 6888-6913 match the forward primer, and positions 7110-7133 match the reverse primer.**

I.2. d. What can you say about the rest of the alignments in the list? **All are shorter, partial matches to one primer or parts of both primers.**

I.3. a. What does the *E-value* of *2e-4* mean? **This denotes a probability of $2x10^{-4}$ or 0.0002.**

I.3. b. Note the names of any significant alignments that have *E-values* less than 0.1. Do they make sense? **There is only one with an *E-value* less than 0.1, and it is from *Arabidoposis* Chromosome 2.**

I.3. e However, the actual length of the fragment includes both ends, so add 1 nucleotide to the result to determine the exact length of the PCR product amplified by the two primers. **246 nucleotides.**

II.2. Describe the source and size of the sequence in which your BLAST hit is located. **The BLAST hit is within a cloned fragment of 123,360 basepairs from *Arabidopsis* Chromosome 2.**

II.3. Identify the feature(s) contained between the nucleotide positions that match the primers, as determined in II.3.d above. What do they tell you about the name of the locus? **The primers overlap with three features that begin at nucleotide positions 6935: a gene and its corresponding mRNA and CDS (coding sequence). The gene is identified as Atg23380, and its product is the curly leaf protein (polycomb group).**

II.4. How many exons are in the *CLF* gene? Where does the reverse primer end? **The *CLF* gene contains 17 exons. The reverse primer ends at position 7133, between the first and second exons.**

II.6. What is this sequence? **The amplicon, or amplified product, expected from the PCR reaction.**

III.6. On what chromosome have you landed? **Chromosome 2, as expected.**

III.9. For each gene diagram, exons (thick bars) are connected by introns (thin lines). Where does the red amplicon match the *CLF* gene? **The first exon and part of the first intron.**

III.10. Click the names of several neighboring genes, and follow links for more information about them. Can you figure out what the prefix At2g means? ***Arabidopsis thaliana* Chromosome 2.**

IV.2. What hits match the *CLF* primer? Do these hits also have matches to the *Ds* primer? What can you conclude? **There is only one match to the *CLF* primer – the same *Arababidopsis* Chromosome 2 clone identified in the first BLAST search. This hit only contains the *CLF* primer, and not the *Ds* primer. So it appears that the *clf-2* mutant sequence is not in the database.**

IV.3. Now examine the hits for the *Ds* primer. Do they make sense? What do they have in common? **Most hits are from corn (*Zea mays*), and many mention *Ds* or the related transposon *Ac*. Several are cloning vectors used in *Ds* mutagenesis.**

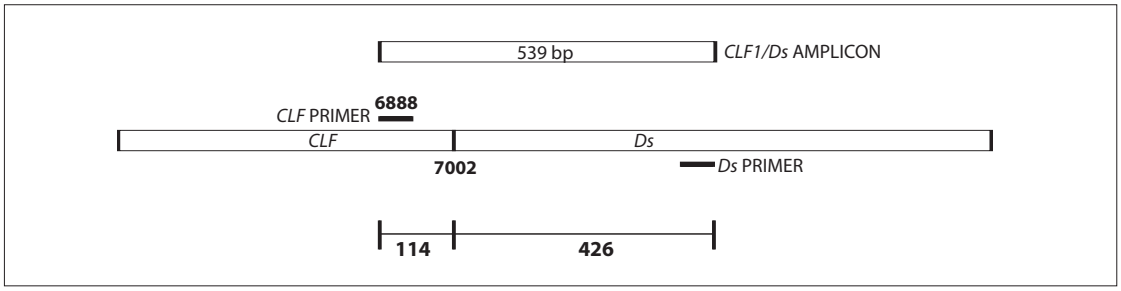
IV.4. Label the following diagram to help you determine the *Ds* insertion site in the *CLF* gene.

a. Scroll to the *Alignments* section and focus on one of the complete *Ds2* sequences. How far into the *Ds2* sequence is the *Ds* primer? **426 bp.**

b. How many *CLF* nucleotides are in the *CLF/Ds* amplicon? **540 bp – 426 bp = 114 bp**.

c. What is the first nucleotide position of the *CLF1* primer in the *CLF* amplicon? **6888.**

d. What is the insertion position of *Ds*? **6888 + 114 = 7002.**



IV.6. Return to the mRNA or CDS entries on the datasheet for the *CLF* gene, to determine which exon or intron the *Ds* transposon is inserted into. **First intron of the *CLF* gene.**

V.4. Follow links to get some quick information about a domain. What do you find? ***CLF* contains a SET domain that encodes a methyltransferase domain (adds methyl groups to other proteins).**

V.6. What can you tell from the graphical overview? **The BLAST hits show strong homology to more than 200 amino acid residues in the second half of the *CLF* protein, which includes the SET domain.**

V.7. What can you conclude from the *E-values*, the titles of the hits, and the *Accession*? **Extremely low *E-values* across a range of organisms – from *Arabidopsis*, *Petunia*, *Oryza* (rice), *Drosophila*, *Mus* (mouse), *Rattus* (rat), and humans – indicate that the SET domain, and other parts of the *CLF* protein, have been strongly conserved through evolutionary time.**

V.8 Follow *Accession* links to sequence datasheets – then on to journal abstracts and articles – to gain insight into how *CLF* and other homeotic proteins control development throughout the plant and animal worlds. **A key insight into the function of *CLF* came from its homology to Enhancer of zeste, a Polycomb-Group protein of *Drosophila*. This is a classic homeotic gene that maintains patterns of gene repression that are set up early in embryonic development. The SET domain exerts its influence by methylating large regions of chromosomes – thus silencing a number of genes in concert.**

CD-ROM CONTENTS

The valuable companion CD-ROM is for exclusive use of purchasers of this DNA Learning Center Kit. To accommodate home or computer lab use by students, all materials may also be reached at the companion Internet site http://bioinformatics.dnalc.org/clf/.

- **Protocol:** a unique online lab notebook with the complete experiment, as well as printable PDF files.
- **Resources:** animations on the discovery of transposable elements, homeotic genes in *Drosophila* development, and key techniques of molecular genetics, from the award-winning Internet sites *DNA from the Beginning* and *DNA Interactive*.

Carolina Biological Supply Company
2700 York Road, Burlington, North Carolina 27215
Phone: 800.334.5551 • Fax: 800.222.7112
Technical Support: 800.227.1150 • www.carolina.com
CB000000000